The Non-Existence of Representative Agents

Matthew O. Jackson* and Leeat Yariv†‡

December 2015

Abstract

We characterize environments in which there exists a representative agent: an agent who inherits the structure of preferences of the population that she represents. The existence of such a representative agent imposes very strong restrictions on individual utility functions – requiring them to be *linear* in the allocation and additively separable in any parameter that characterizes agents (e.g., a risk aversion parameter, a discount factor, etc.). In particular, commonly used classes of utility functions (exponentially discounted utility functions, CRRA or CARA utility functions, logarithmic functions, etc.) do not admit a representative agent.

JEL Classification Numbers: D72, D71, D03, D11, E24

Keywords: Representative Agents, Collective Decisions

1 Introduction

Properties of individual behavior do not generally carry over when aggregated. For example, the classic Sonnenschein-Mantel-Debreu Theorem (Sonnenschein, 1973; Mantel, 1974; Debreu, 1974) illustrated that under a set of standard assumptions on individual demand, there are essentially no restrictions on aggregate demand.

The literature has often taken the approach of assuming a *representative agent*, one whose choices or preferences mirror those aggregated across society. The notion itself can be traced back to Edgeworth (1881) and Marshall (1890), but was theoretically founded only in the mid-twentieth century. Gorman (1953) illustrated that if indirect utility functions take a particular form, termed the 'Gorman Form,' then a researcher can treat a society of utility maximizers *as if* it consists of one 'representative' agent.

*Department of Economics, Stanford University, the Santa Fe Institute, and CIFAR. http://www.stanford.edu/~jacksonm e-mail: jacksonm@stanford.edu

†Division of the Humanities and Social Sciences, Caltech. http://www.hss.caltech.edu/~lyariv/index.htm e-mail: lyariv@hss.caltech.edu

‡We are grateful to William Brainard for very helpful suggestions.

Since the publication of the Lucas Critique (1976), micro-founding economic models to the level of individual behavior has become pervasive. Given the challenges of handling the details of heterogeneous societies, the use of a representative agent as a modeling tool is standard practice. Often, however, practitioners effectively assume much more than the results of Gorman (1953) guarantee. Indeed, researchers frequently impose a particular structure on individual utility functions, commonly identified by one or more parameters that could vary in the population as a whole (e.g., exponential discounting with different discount rates, CRRA or CARA utility functions with different risk-aversion parameters, etc.). The representative agent is then assumed to be characterized by preferences from the same class. We identify environments where such an assumption is plausible.

We fully characterize the classes of preferences for which representative agents exist, identifying the conditions under which *the population's preferences can be represented by an agent who has preferences in the same class*. As we show, the existence of such a representative agent imposes strong restrictions on individual utility functions. In particular, the commonly used classes of utility functions (exponentially discounted utility functions, CRRA or CARA utility functions, logarithmic, mean-variance, concave functions, etc.) cannot be aggregated to generate a representative agent who is characterized by preferences from the same class. When agents are evaluating private allocations, only utility functions that are linear in the allocation and additively separable in agents' parameters admit representative agents. Since effectively none of the literature using representative agents assumes a linear utility function, this means that none of those models of representative agents really represents a heterogeneous society with preferences from the assumed class.

Our results indicate the perils of using representative agent models. The behaviors that such models predict will generally not reflect aggregate behavior in society, unless agents exhibit very particular preferences, ones that are rarely assumed. Furthermore, much of the literature that uses representative agents defines optimal policies as those that maximize the representative agent's welfare. Our results show that such policies may not, in fact, be welfare maximizing for the underlying population. Thus, representative agent analyses cannot be sure to properly evaluate first-best, or second-best policies, unless the agents in the underlying population have linear preferences. Moreover, the behavior of utility maximizing agents may not be properly reflected by the representative agent's utility-maximizing behavior.

The insights of this paper are in the spirit of Jackson and Yariv (2015), where we showed that there is no utilitarian aggregation of exponentially-discounted preferences that satisfies time consistency. Here we show that such impossibilities are a much more pervasive phenomenon – applying to many different preference formulations and for quite general sources of heterogeneity, and any weightings – and can be argued quite directly.

2 Representative Agents with Private Allocations

We first consider the case in which individuals each have their own allocation and the representative agent evaluates the aggregate/average allocation. For example, the allocation could be consumption, investment, and/or savings levels. Agents may then have different discount factors, risk aversion parameters, or other heterogeneity in preferences, which may depend on their age, wealth levels, idiosyncrasies, or other features. A representative agent then evaluates the representative allocation.

2.1 Definitions – Private Allocations

Agents evaluate alternatives, such as consumption profiles, investments, savings, etc. An alternative is denoted by x and lies in a set D_x, which is either an open or compact subset of $\mathbb{R}$.

The heterogeneity of agents – for instance in their risk aversion parameters, discount factors, etc. – is indexed by $a \in D_a$, where $D_a \subset \mathbb{R}$ is compact. We normalize the minimum value of a to be 0 and the maximum to be 1, without loss of generality.

Utility functions are analytic functions of the form $V : D_x \times D_a \to \mathbb{R}$. Utilities in the society are denoted by $\Delta = \{V(x; a)\}$. Utility functions are such there exists at least one $x^* \in D_x$ for which $V(x^*; a)$ is monotone (increasing or decreasing) in a.

Although we consider a case in which x and a are each one-dimensional for ease of notation, the proofs extend directly to the multi-dimensional case (without changes, using a re-interpretation of notation[1]).

We say that there exists *a representative agent* if for any $\lambda_1, ..., \lambda_n \geq 0, \sum_{i=1}^{n} \lambda_i = 1$, and $(a_1, \ldots, a_n) \in D_a^n$, there exists $\bar{a} \in D_a$ such that for all $(x_1, \ldots, x_n) \in D_x^n$:

$$\sum_{i=1}^{n} \lambda_i V(x_i; a_i) = V(\sum_{i=1}^{n} \lambda_i x_i; \bar{a}).$$

Given that the set of Pareto optimal allocations are generally characterized as those that maximize $\sum_{i=1}^{n} \lambda_i V(x_i; a_i)$ for some weights, this is a minimal requirement to ensure that the representative agent properly characterizes the Pareto frontier, or any other welfare evaluation that is made using some weights on the preferences of the population's agents.

In our formulation, $\bar{a}$ is the representative agent's preference parameter. The representative agent's utility function is often assumed to take the form $bexp(c - \bar{a}x)$, $(c + bx)^{\bar{a}}/\bar{a}$, $\bar{a}\ log(x)$, etc.[2] Is it possible to have such preferences represent an economy in which all

[1] That is, in the representation of analytic functions, interpret $(x - y)^k$ to be the appropriate multi-dimensional analog and the proofs go through as stated.

[2] In these forumlations, b and c are taken as constants. For instance, the form $b \exp(c - \bar{a}x)$ with $b = 1$ and $c = 0$ would correspond to a representative agent with a CARA utility function and the form $(c+bx)^{\bar{a}}/\bar{a}$ with $c = 0$ and $b = 1$ would correspond to a representative agent with a CRRA utility function.

agent have a utility function from one of these classes (with heterogeneous parameters)? The answer is no.

The existence of a representative agent is a sort of convexity requirement on the space of utility functions. For example, in the special case in which the x_i's are all equal, this requires that a convex combination of the utility functions of the agents is also an admissible utility function.

We also consider a stronger notion of representation, which turns out to be useful as a step in our proofs (but is not assumed in the main theorems) and is also of independent interest. We say that there exists a *strongly representative agent* if for any $\lambda_1, ..., \lambda_n \geq 0, \sum_{i=1}^{n} \lambda_i = 1$, $(a_1, \ldots, a_n) \in D_a^n$, and $(x_1, \ldots, x_n) \in D_x^n$:

$$\sum_{i=1}^{n} \lambda_i V(x_i; a_i) = V\left(\sum_{i=1}^{n} \lambda_i x_i; \sum_{i=1}^{n} \lambda_i a_i\right). \tag{1}$$

2.2 Classes of Utility Functions that Admit Representative Agents

The following is the characterization of utility functions that admit the existence of a representative agent.

Theorem 1 *There exists a representative agent if and only if there exists an analytic and monotone function $h(a)$ and a constant b such that $V(x; a) = h(a) + bx$ for all x, a.*

The proofs of our results appear in Section 5.

Note that the structure characterized by Theorem 1 is not satisfied by utility functions that are commonly used in economic modeling. For example, CRRA or CARA utility functions do not satisfy the restriction, nor do strictly concave utility functions. In such cases, the theorem suggests that assuming a representative agent with the same "type" of preferences would generate inaccurate predictions of aggregate preferences and policy recommendations that are not necessarily welfare enhancing.

The existence of a strongly representative agent imposes even harsher restrictions.

Proposition 1 *There exists a strongly representative agent if and only if there exist constants b_1, b_2, and c such that $V(x; a) = b_1 a + b_2 x + c$ for all x, a.*

Proposition 1 states that a strongly representative agent exists only when utility functions are additively separable in the preference parameter and the allocation, and *linear* in both. The intuition is that if a strongly representative agent exists, a marginal change in either the alternative or the parameter of any individual has a proportional effect on the representative agent's alternative and utility parameter (where the proportional factor corresponds to the individual's weight in society). This implies the linear structure of utility functions.

3 Representative Agents with Common Allocations

The case of private allocations relates to most of the work utilizing representative agents in macroeconomics and finance. However, common consumption is often assumed in models falling under the umbrella of political economy and public finance. In these models, agents do not consume different allocations but instead consume some public good. As we now show, the existence of a representative agent in such environments is still very restrictive, but entails a different sort of separability.

3.1 Definitions – Common Allocations

Under a common allocation, all the x_i's are restricted to be equal to a common x and the above definitions simplify.

We say that there exists *a representative agent* if for any $\lambda_1, ..., \lambda_n \geq 0, \sum_{i=1}^{n} \lambda_i = 1$, and for any $a_1, ..., a_n \in [0,1]$, there exists $\bar{a}$ such that :

$$\sum_{i=1}^{n} \lambda_i V(x; a_i) = V(x; \bar{a})$$

for all x.

Again, the stronger notion of a strongly representative agent imposes more structure on the parameter identifying the representative agent's preferences. We say that there exists *a strongly representative agent* if for any $\lambda_1, ..., \lambda_n \geq 0, \sum_{i=1}^{n} \lambda_i = 1$, and for any $a_1, ..., a_n \in [0,1]$,

$$\sum_{i=1}^{n} \lambda_i V(x; a_i) = V\left(x; \sum_{i=1}^{n} \lambda_i a_i\right). \tag{2}$$

3.2 Restrictions for Representative Agents with Common Allocations

We first provide a characterization of utility functions for which a representative agent exists.

Theorem 2 *There exists a representative agent if and only if there exist analytic functions $h(a), f(x), g(x)$ such that $V(x; a) = h(a) f(x) + g(x)$ for all x, a, where $h(\cdot)$ is monotone and $f(x^*) \neq 0$.*

While the restrictions for *purely* public consumption are weaker than those for private consumption, they are still sufficiently strong as to rule out nearly all of the commonly assumed utility functions. From the examples mentioned so far, CARA and CRRA utility functions with risk-aversion parameters do not satisfy the restrictions of Theorem 2, nor do concave loss functions with bliss points serving as parameters – e.g., single-peaked preferences.

The following proposition addresses the existence of a strongly representative agent when allocations are common.

Proposition 2 *There exists a strongly representative agent if and only if there exist analytic functions $f(x), g(x)$ such that $V(x; a) = af(x) + g(x)$ for all x, a.*

The intuition behind Proposition 2 is similar to that underlying the intuition of Proposition 1. When an average representative agent exists, a marginal change in one individual's utility parameter has a proportional impact on the marginal change of the representative agent's utility parameter. This maps into a linearity requirement with respect to the utility parameter a.

4 Discussion

The assumption of a representative agent, who has preferences representing the aggregate of the population, is commonplace in modern Economics. We have shown that for the representative agent to inherit the structure of preferences in the population she represents, harsh restrictions need to be satisfied. In particular, parametrized utility functions need to exhibit separability with respect to terms pertaining to parameters and those pertaining to alternatives. Unfortunately, these restrictions are not satisfied by commonly used classes of utility functions. For instance, a society in which each agent is characterized by a CRRA or CARA utility, or a concave loss function, etc., does not admit a representative agent with a similar utility function.

While others have pointed out the challenges of using a representative-agent model when individuals in society interact strategically (see Kirman, 1992 and references therein), the results in this paper are even more fundamental. They illustrate an impossibility result for many classes of utility functions, including practically all those studied in the literature. The paper calls for a more careful consideration of the type of preferences representative agents can have.

5 Proofs

We first prove Proposition 1, which is useful for the proof of Theorem 1.[3]

Proof of Proposition 1: We show that (1) implies that there exist constants b_1, b_2, and c such that $V(x; a) = b_1 a + b_2 x + c$ for all x, a, as the converse is obvious.

[3] We note that the two propositions do not require any monotonicity of V in their proofs. Monotonicity is only used in the proof of the Theorems.

Since $V(x,a)$ is analytic, for any x', a' there exist a set of scalars $(c_{jk})_{j,k\in 0,\ldots\infty}$ for which

$$V(x,a) = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} c_{jk}(x-x')^j(a-a')^k, \tag{3}$$

for any x, a.

Using (3) and noting that $\sum_i \lambda_i = 1$ we can write (1) as

$$\begin{aligned}
&\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{i=1}^{n} \lambda_i c_{jk}[(x_i-x')^j(a_i-a')^k] \\
= &\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} c_{jk}[\sum_{i=1}^{n}\lambda_i(x_i-x')]^j[\sum_{i=1}^{n}\lambda_i(a_i-a')]^k.
\end{aligned}$$

Noting that the terms on the two sides are the same whenever $j+k\leq 1$, this implies that

$$\begin{aligned}
&\sum_{j,k:j+k>1}\sum_{i=1}^{n} \lambda_i c_{jk}[(x_i-x')^j(a_i-a')^k] \\
= &\sum_{j,k:j+k>1} c_{jk}[\sum_{i=1}^{n}\lambda_i(x_i-x')]^j[\sum_{i=1}^{n}\lambda_i(a_i-a')]^k.
\end{aligned}$$

Thus, for admissible λ_i's and a_i's it must be that:

$$\sum_{j,k:j+k>1} c_{jk}\left(\left[\sum_{i=1}^{n}\lambda_i[(x_i-x')^j(a_i-a')^k]\right] - \left[[\sum_{i=1}^{n}\lambda_i(x_i-x')]^j[\sum_{i=1}^{n}\lambda_i(a_i-a')]^k\right]\right) = 0. \tag{4}$$

Since x', a' were arbitrary, take $a' \in (0,1)$ and $x' < \sup D_x$. Consider $a_i > a'$ and set $x_i = \hat{x} > x'$ for all i so that $\sum_{i=1}^{n}\lambda_i(\hat{x}-x') = (\hat{x}-x')$. In this case, the term corresponding to $j=k=1$ vanishes. For any $k\geq 2$ or $j\geq 2$, by Jensen's inequality,

$$\left(\left[\sum_{i=1}^{n}\lambda_i(a_i-a')^k\right] - \left[\sum_{i=1}^{n}\lambda_i(a_i-a')\right]^k\right) > 0. \tag{5}$$

Multiplying both terms by $(\hat{x}-x')^j$, and noting that

$$(\hat{x}-x')^j = \sum_i \lambda_i(\hat{x}-x')^j$$

allows us to rewrite (5) as

$$\left(\sum_{i=1}^{n}\lambda_i(\hat{x}-x')^j(a_i-a')^k - \left[\sum_{i=1}^{n}\lambda_i(\hat{x}-x')\right]^j\left[\sum_{i=1}^{n}\lambda_i(a_i-a')\right]^k\right) > 0$$

for any $k \geq 2$ or $j \geq 2$. Thus, to satisfy (4) it must be the case that $c_{jk} = 0$ for all $j \geq 0, k \geq 2$, and $j \geq 2, k \geq 0$. Therefore, equality 4 boils down to:

$$c_{11}\left(\left[\sum_{i=1}^{n} \lambda_i[(x_i - x')(a_i - a')]\right] - \left[[\sum_{i=1}^{n} \lambda_i(x_i - x')][\sum_{i=1}^{n} \lambda_i(a_i - a')]\right]\right) = 0.$$

In particular, choose $\lambda_1 = \lambda_2 = 1/2$, $x_1 = x', x_2 = x' + \varepsilon, a_1 = a'$,and $a_2 = a' + \varepsilon$ for sufficiently small $\varepsilon > 0$ so that $x_2 \in D_x$ and $a_2 \in D_a$. Then, we get:

$$c_{11}\left(\frac{\varepsilon^2}{2} - \frac{\varepsilon}{2} \cdot \frac{\varepsilon}{2}\right) = c_{11}\frac{\varepsilon^2}{4} = 0.$$

It follows that $c_{11} = 0$ and we can rewrite (3) as:

$$V(x, a) = c_{00} + c_{01}(x - x') + c_{10}(a - a'),$$

for all x, a, which implies the claim. ∎

Proof of Theorem 1: Define

$$\tilde{h}(a) \equiv V(x^*; a).$$

Note that $\tilde{h}(\cdot)$ is analytic and monotone in a and therefore $\tilde{h}^{-1} : [0, 1] \to D_a$ is analytic and monotone as well. Let

$$G(x; a) = V(x; \tilde{h}^{-1}(a)).$$

Note that for any $\lambda_1, ..., \lambda_n \geq 0, \sum_{i=1}^{n} \lambda_i = 1$, and for any $a_1, ..., a_n \in D_a^n$, there exists $\bar{a}$ such that for all $x_1, ..., x_n$,

$$\sum_{i=1}^{n} \lambda_i G(x_i; a_i) = G(\sum_{i=1}^{n} \lambda_i x_i; \bar{a}),$$

Since V satisfies this property. In particular, by choosing $x_1 = x_2 = ... = x_n = x^*$,

$$\sum_{i=1}^{n} \lambda_i G(x^*; a_i) = \sum_{i=1}^{n} \lambda_i a_i = G(x^*; \bar{a}) = \bar{a}.$$

Therefore, G satisfies the assumptions of Proposition 1, so that there exist constantsb_1, b_2, c such that

$$G(x; a) = b_1 a + b_2 x + c,$$

which, when setting $h(a) = b_1\tilde{h}(a) + c$ and $b = b_2$, implies that

$$V(x; a) = h(a) + bx.$$

This establishes the theorem, as the converse is direct. ∎

Proof of Proposition 2: We show that (2) implies that there exist analytic functions $f(x), g(x)$ such that $V(x;a) = af(x) + g(x)$ for all x, a, as the converse is obvious.

As before, we can use the formulation in (3): for any x', a' there exist a set of scalars $(c_{jk})_{j,k\in 0,\ldots\infty}$ for which

$$V(x,a) = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} c_{jk}(x-x')^j(a-a')^k.$$

Therefore, noting that $\sum_i \lambda_i = 1$ we can write (2) as

$$\begin{aligned} & \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} c_{jk}(x-x')^j[\sum_{i=1}^{n}\lambda_i(a_i-a')^k] \\ = \;& \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} c_{jk}(x-x')^j[\sum_{i=1}^{n}\lambda_i(a_i-a')]^k. \end{aligned}$$

The terms on the two sides are the same whenever $k=0$ and $k=1$, hence we have

$$\begin{aligned} & \sum_{j=0}^{\infty}\sum_{k=2}^{\infty} c_{jk}(x-x')^j[\sum_{i=1}^{n}\lambda_i(a_i-a')^k] \\ = \;& \sum_{j=0}^{\infty}\sum_{k=2}^{\infty} c_{jk}(x-x')^j[\sum_{i=1}^{n}\lambda_i(a_i-a')]^k. \end{aligned}$$

Thus, for admissible λ_i's and a_i's it must be that we can restrict attention to $k \geq 2$:

$$\sum_{j=0}^{\infty} c_{jk}(x-x')^j \sum_{k=2}^{\infty}\left([\sum_{i=1}^{n}\lambda_i(a_i-a')^k] - [\sum_{i=1}^{n}\lambda_i(a_i-a')]^k\right) = 0. \tag{6}$$

Since x', a' was arbitrary, take $a' \in (0,1)$. Then this must hold for any given $a_i > a'$ and $x > x'$. Note that by Jensen's inequality

$$\left([\sum_i \lambda_i(a_i-a')^k] - [\sum_i \lambda_i(a_i-a')]^k\right) > 0$$

for any $a_i > a'$ and all $k \geq 2$. Thus, given $x > x'$, from the above it is only possible to satisfy (6) if $c_{jk} = 0$ for all $j \geq 0$ and all $k \geq 2$.

Therefore, we can rewrite (3) as:

$$V(x,a) = \sum_{j=0}^{\infty}\left(c_{j0} + c_{j1}(a-a')\right)(x-x')^j,$$

for all x, a, which implies the claim. ▮

Proof of Theorem 2: Define

$$h(a) \equiv V(x^*;a).$$

Notice that $h(\cdot)$ is analytic and monotone in a and therefore $h^{-1}:[0,1]\to D_a$ is analytic and monotone as well. Now let

$$G(x;a)=V(x;h^{-1}(a)).$$

Note that for any $\lambda_1,...,\lambda_n\geq 0, \sum_{i=1}^{n}\lambda_i=1$, and for any $a_1,...,a_n\in D_a^n$, there exists $\bar{a}$ such that for all x,

$$\sum_{i=1}^{n}\lambda_i G(x;a_i)=G(x;\bar{a}),$$

Since V satisfies this property. In particular:

$$\sum_{i=1}^{n}\lambda_i G(x^*;a_i)=\sum_{i=1}^{n}\lambda_i a_i=G(x^*;\bar{a})=\bar{a}.$$

Therefore, G satisfies the assumptions of Claim 1, so that there exist analytic functions $f(x), g(x)$ such that

$$G(x;a)=af(x)+g(x),$$

which, in turn, implies that

$$V(x;a)=h(a)f(x)+g(x).$$

This establishes the theorem, as the converse is direct. ▮

6 References

Debreu, Gerard (1974), "Excess-demand Functions," *Journal of Mathematical Economics*, 1, 15-21.

Edgeworth, Francis Y. (1881), *Mathematical Psychics*, London: Kegan Paul.

Gorman, William M. (1953), "Community Preference Fields," *Econometrica*, 21(1), 63-80.

Jackson, Matthew O. and Leeat Yariv (2015), "Collective Dynamic Choice: The Necessity of Time Inconsistency," *American Economic Journal: Microeconomics*, 7(4): 150-178.

Kirman, Alan P. (1992), "Whom or What does the Representative Agent Represent," *Journal of Economic Perspectives*, 6(2), 117-136.

Lucas, Robert E. (1976), "Econometric Policy Evaluation: A Critique," in K. Brunner and A. H. Meltzer (eds.), *The Phillips Curve and Labor Markets*, Vol. 1 of Carnegie-Rochester Conference Series on Public Policy, Amsterdam: North-Holland.

Mantel, Rolf R. (1974), "On the Characterization of Aggregate Excess-demand," *Journal of Economic Theory* , 7, 348-353.

Marshall, Alfred (1890), Principles of Economics, London: Macmillan.

Sonnenschein, Hugo (1973), "Do Walras' Identity and Continuity Characterize the Class of Community Excess-demand Functions?," *Journal of Economic Theory*, 6, 345-354.